Optimal unemployment insurance with monitoring and sanctions

Boone, J.; Fredriksson, P.; Holmlund, B.; van Ours, J.C.

Published in:
Economic Journal

Publication date:
2007

Optimal Unemployment Insurance with Monitoring and Sanctions*

Jan Boone[†] Peter Fredriksson[‡] Bertil Holmlund[§] Jan van Ours[¶]

Revised version: 7 April 2004

Abstract

This paper analyzes the design of optimal unemployment insurance in a search equilibrium framework where search effort among the unemployed is not perfectly observable. We examine to what extent the optimal policy involves monitoring of search effort and benefit sanctions if observed search is deemed insufficient. We find that introducing monitoring and sanctions represents a welfare improvement for reasonable estimates of monitoring costs; this conclusion holds both relative to a system featuring indefinite payments of benefits and a system with a time limit on unemployment benefit receipt.

JEL-classification: J64, J65, J68

Keywords: Unemployment insurance, search, sanctions

*We gratefully acknowledge comments from an anonymous referee, David Grubb, Åsa Rosén and seminar participants at IZA, Bonn, Stockholm University, University of Copenhagen, University of Gothenburg, and University of Cyprus. Fredriksson and Holmlund acknowledges the the financial support from the Swedish Council for Working Life and Social Research. Boone acknowledges the financial support from NWO, KNAW and VSNU through a Vernieuwingsimpuls grant.

†Department of Economics and CentER for Economic Research, Tilburg University, OSA and CEPR. Postal address: Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: j.boone@uvt.nl. Boone gratefully .

‡Department of Economics, Uppsala University and IFAU. Postal address: Box 513, SE-751 20 Uppsala, Sweden. E-mail: peter.fredriksson@nek.uu.se.

§Department of Economics, Uppsala University and CESifo. Postal address: Box 513, SE-751 20 Uppsala, Sweden. E-mail: bertil.holmlund@nek.uu.se.

¶CentER for Economic Research, Tilburg University, OSA, IZA and CEPR. Postal address: Box 90153, 5000 LE Tilburg, The Netherlands. E-mail vanours@uvt.nl.

1 Introduction

It is generally accepted that public provision of unemployment insurance (UI) is socially desirable in a world with risk averse individuals. However, it is also well established that the provision of UI does not come without adverse incentive effects. For example, more generous UI benefits is likely to reduce search effort and raise wage pressure, thus causing some increase in unemployment. The problem facing policy makers is thus to strike an optimal balance between the insurance benefits on the one hand, and the adverse incentive effects on the other hand. This problem has been the subject of several recent papers. Our paper contributes to this literature by recognizing that the government may condition benefit payments on (imperfectly) observed search effort. This leads us to an analysis of optimal UI design in a search equilibrium framework where the government has several policy instruments at its disposal, including the benefit level, the rate at which search effort is monitored, and the magnitude of the sanction in case search effort is regarded as insufficient. We find that a system with monitoring and sanctions represents a welfare improvement relative to other alternatives for reasonable estimates of the monitoring costs. In particular, the monitoring and sanction system leads to higher welfare than a system with time limits.

Our results on the desirability of monitoring can be contrasted with a well-known result that dates back to Becker's (1968) celebrated paper on optimal crime deterrence. In Becker's analysis (as in ours), monitoring is costly because resources have to be spent on detecting crime (violations of search requirements). Punishment, in the form of a fine (sanction), goes without cost since it involves a transfer of money from one individual to others. To deter crime the expected fine, i.e., the probability of being caught times the fine, should be big enough. By raising the fine, monitoring costs can be reduced without affecting incentives for crime. However, Becker's analysis presupposes risk neutral agents. When agents are risk averse and there are errors in the monitoring technology, Becker's result need not hold. If the monitoring technology is plagued by Type II errors, some complying individuals are sanctioned and these individuals will be subjected to substantial welfare losses when fines are high.[1]

[1] This squares with the conclusion in Polinsky and Shavell (1979). They conclude that risk aversion weakens the case for the Beckerian policy prescription. Furthermore, they note that the possibility of making Type II errors reinforces this conclusion. See Garoupa (1997) and Polinsky and Shavell (2000) for recent surveys of the economic theory of law

Shavell and Weiss (1979) presented a seminal analysis of optimal sequencing of benefit payments over the spell of unemployment. The key result was that the benefit level should decline monotonically over the unemployment spell, because such a profile involves stronger incentives to search. Recently a number of papers have extended the analysis of Shavell and Weiss. One strand of the literature adds additional policy instruments; Hopenhayn and Nicolini (1997) is a case in point. Another strand of the literature (e.g. Cahuc and Lehmann, 2000, and Fredriksson and Holmlund, 2001) takes account of firm behavior and allows for endogenous wage determination. Endogenous wages is potentially important since a declining benefit profile can raise wage pressure. Wage pressure may rise because it is the value of unemployment upon unemployment entry that enters the worker's outside option. The analysis in Fredriksson and Holmlund (2001), however, suggests that there is still a case for having a declining profile of benefit payments.

The contributions reviewed above, and most of the other literature on optimal UI, do not consider that the government can make the receipt of benefits dependent on the unemployed worker's search effort. As documented by Grubb (2001), existing UI systems condition benefit payments on performance criteria such as "availability for work" and "active job search". These criteria are enforced by some degree of monitoring of the benefit claimants. The requirements for job search show substantial variations across countries.

Failure to meet search requirements may result in a benefit sanction, i.e., a temporary or permanent cut in benefits. A typical duration of sanctions for a first refusal of a suitable job offer is two to three months. Observed sanction rates – the total number of sanctions over a year relative to the stock of beneficiaries – also vary substantially across countries. For example, sanctions due to insufficient search hovered around 30 percent in the United States in the late 1990s, whereas other countries (Germany, Denmark, Norway) appear to have undertaken no sanctions related to search inactivity; see Grubb (2001) for further details.

Recent empirical work has shed light on the effects of changes in search requirements and monitoring of job search. The arguably most convincing evidence is based on randomized experiments undertaken in the United States. The "treatments" in these experiments involved the number of employer contacts, the required documentation and the frequency of verification.[2] These

enforcement.

[2]See OECD (2001), Johnson and Klepinger (1994), Benus et al (1997) and Black et al

studies indicate that either more intensive monitoring or more demanding search requirements tend to reduce the length of benefit claims. Recent non-experimental evidence from the Netherlands and Switzerland also suggest that the imposition of sanctions substantially raises the transition rate to employment (Abbring et al., 1997; Lalive et al., 2002; van den Berg et al., 2004). Our reading of the bulk of the evidence is that more intensive monitoring and more stringent search requirements do matter for search activity and transitions out of unemployment.[3]

The literature on monitoring and sanctions in the context of UI is very small. The study most closely related to what we do in the present paper is Boone and van Ours (2000). The model is a version of the Pissarides (1990) search and matching model and has similarities with the model in Fredriksson and Holmlund (2001). A key feature of the model is that the unemployed and insured worker can affect the probability of continued UI receipt by the choice of search effort; the higher the search effort, the lower the risk of being exposed to a benefit sanction.

The analysis of monitoring and sanctions is clearly related to the analysis of the optimal sequencing of UI benefits. Indeed, one can think of the declining profile of benefit payments as an indirect "sanction" on deficient job search. The defining characteristic of a monitoring and sanction system, however, is that the risk of being sanctioned depends *directly* on search activity. This feature can have substantial implications for policy prescriptions. Let us illustrate this point by considering a world with risk aversion and a finite arrival rate of job offers. In this situation, a system with a time limit on UI benefit receipt can never have "Beckerian properties". The reason is that some workers will be penalized as time passes. However, a Becker-type solution is a distinct possibility when the risk of being penalized depends directly on search. If the monitoring technology is perfect, the government can implement the optimal search intensity by threatening to impose the maximal sanction.[4] A disadvantage of the monitoring and sanctions system is that resources are needed for monitoring. The argument in favor of mon-

(2003).

[3]There is at least one study, van den Berg and van der Klaauw (2001), that fails to confirm that more intensive monitoring affects transitions out of unemployment. The authors conjecture that the result may reflect that more stringent monitoring of formal search induces a substitution away from informal search channels.

[4]This a viable strategy with risk aversion since nobody will be sanctioned in equilibrium with a perfect monitoring technology.

itoring and sanctions depends crucially on the costs associated with such a system. One of the contributions of the paper is to derive an upper bound on the marginal cost of monitoring where the monitoring system dominates the pure time limits for all marginal costs below this upper bound.

In this paper we extend the contributions by Boone and van Ours (2000) and Fredriksson and Holmlund (2001) by offering a normative analysis of a benefit system with costly monitoring and sanctions. The basic model features two benefit levels which can be thought of as unemployment insurance (UI) and unemployment assistance (UA), respectively. Workers who receive UI are monitored at a certain rate and, with some probability, exposed to a benefit sanction. The probability of being sanctioned depends on the worker's search effort and the precision at which search effort can be observed by the UI provider. Sanctioned workers receive UA, they are not monitored, and they need to become reemployed before they are entitled to UI. We are concerned with the characteristics of the optimal benefit system when there are four available policy instruments: the level of benefits in UI and UA (the difference between the two representing the sanction), the rate at which the unemployed worker entitled to UI is monitored, and the precision of the monitoring technology that determines how the agent's search effort affect the probability of a sanction.

The next section of the paper presents the basic model. Section 3 derives some analytical results concerning the properties of the optimal benefit system. In section 4, we turn to a numerical analysis of the optimal benefit system. Section 5 concludes.

2 The Model

2.1 The Labor Market

We consider an economy with a fixed labor force, which is normalized to unity. Workers are either employed or unemployed and have infinite horizons. Time is continuous. An employed worker is separated from his job at an exogenous Poisson rate ϕ. Upon entering unemployment, the worker is immediately eligible for UI benefits.

Recipients of UI benefits are monitored with respect to their search behavior. If they fail to meet certain search requirements, they are exposed to a benefit withdrawal (a sanction). We assume that the sanction lasts for the remainder of the unemployment spell. At every instant, there are thus two groups of unemployed workers: *eligible* workers who receive benefits and *sanctioned* workers who have been exposed to a benefit withdrawal.

Let α^j, $j = e, s$ denote the exit rate from unemployment to employment for an $e$ligible and a $s$anctioned worker, respectively. The exit rates differ between the two groups to the extent that their search effort differ. Let s^j, $j = e, s$, denote search effort. The effective number of searchers in the economy is then given as $S = s^e u^e + s^s u^s$, where u^j is the number of unemployed in category j.

The matching function is of the usual constant returns to scale variety: $H = H(S, v)$, where v is the number of vacancies. Let $\theta \equiv v/S$ denote labor market tightness. The probability per unit time that individual i escapes unemployment state j is then obtained as $\alpha_i^j \equiv s_i^j H(S, v)/S = s_i^j \alpha(\theta)$. Also, $\alpha(\theta) = H(S, v)/S = H(1, \theta)$ and hence $\alpha'(\theta) > 0$; the tighter the labor market, the easier to find a job. Firms fill vacancies at the rate $q(\theta) = H(S, v)/v = H(1/\theta, 1)$, and thus $q'(\theta) < 0$; the tighter the labor market, the more difficult to fill a vacancy. By constant returns to scale, we also have $\alpha(\theta) = \theta q(\theta)$.

While unemployed and receiving UI benefits, an unemployed agent is monitored at rate μ. We think of monitoring as random inspections of the worker's search activity. Given monitoring, there is some probability that the observed search effort does not meet the search requirement, in which case the worker is sanctioned. Let $\pi(s^e)$ denote the probability of being sanctioned upon inspection of search effort, implying that UI recipients loose entitlement at the rate $\mu\pi(s^e)$.

Having defined the relevant transition rates, we can formulate the steady state flow equilibrium relationships of the labor market:

$$\phi n = \alpha^e u^e + \alpha^s u^s \tag{1}$$

$$\alpha^s u^s = \mu \pi u^e \tag{2}$$

where $n = 1 - u^e - u^s$ denotes total employment in the economy. The first equation pertains to employment whereas the second equation pertains to the state of unemployment with a sanction. Now we can use (1) and (2) to solve for employment:

$$n = \frac{\lambda (\alpha^e + \mu\pi)}{\phi + \lambda(\alpha^e + \mu\pi)} \tag{3}$$

where $\lambda \equiv u^e/(u^e + u^s) = \alpha^s/(\alpha^s + \mu\pi))$ is the ratio of eligible unemployment to total unemployment.

2.2 Monitoring and Sanctions

Let us make the monitoring and sanctions technology explicit. We choose a reduced form specification which allows us to have as special cases indefinite payments of UI benefits ($\mu = 0$), finite duration of UI benefit receipt ($\mu > 0$ and $\pi(s_i^e) = 1$), and a monitoring and sanctions technology. In particular, we assume that the probability of being sanctioned upon inspection depends linearly on search: $\pi(s_i^e) = 1 - \sigma s_i^e$. Proposition 2 below gives conditions under which $\sigma > 0$ is optimal. Further, we require that $\pi\left(s_i^e\right) \geq 0$ for all $s_i^e \in [0, 1]$, which, in turn, implies that $\sigma \in [0, 1]$.

The parameter σ measures to which extent the sanction probability depends on an agent's own search effort. One way to interpret σ is that it indexes the precision of the inspection technology. For instance, $\sigma = 0$ corresponds to the situation where it is determined by lottery if the agent has searched to rule or not; therefore everyone who is monitored is sanctioned irrespective of search intensity. Alternatively, $\sigma = 0$ can be seen as a UI system with a time limit, as in Fredriksson and Holmlund (2001). If, on the other hand, σ is strictly positive the agent's search effort matters for the sanction probability. The higher is σ, the higher the precision with which an agent's search effort is observed and rewarded.[5]

Whereas $\sigma = 0$ gives little direct incentive to search, it is an inexpensive system to operate. This is due to the fact that there are no inspections of agents' search effort. On the other hand, $\sigma > 0$ gives a direct incentive to search but also implies that more monitoring officials are needed in order to inspect agents' search intensities. So the monitoring cost per monitored agent is increasing in σ.

More precisely, we assume that the cost of running the monitoring and sanctioning system, C, is given by:

$$C = c\left(\sigma\right)\mu u^e w \tag{4}$$

The cost of running the UI-system is increasing in the number of monitored individuals (μu^e). The rate of increase is determined by $c\left(\sigma\right) \geq 0$. This

[5] From a more general point of view, it is possible to derive this technology from "first principles" with the aid of a few assumptions. We present this derivation in Appendix A.

cost depends on the precision of the inspection technology with $c'(\sigma) \geq 0$ and $c(0) = 0$. We think of the inspection of search as a labor intensive activity and, therefore, the monitoring cost is proportional to the aggregate wage w.

2.3 Worker Behavior

The employed worker's (indirect) instantaneous utility is determined by his wage, w. The unemployed worker receives unemployment benefits, B, as long as he is eligible. When sanctioned, he receives Z. We show in proposition 1 below that $B > Z$. We assume that workers do not have access to a capital market, so consumption equals income at each instant.

We take the utility functions to be strictly concave in income and leisure. The unemployed worker's instantaneous utility is decreasing in search effort, since search reduces time available for leisure. The utility function for the eligible unemployed worker is $\upsilon(B, s_i^e)$ and for the sanctioned worker it is $\upsilon(Z, s_i^s)$. The employed worker's utility is given by $\upsilon(w_i, \overline{h})$, where $\overline{h}$ denotes hours of work; we take $\overline{h}$ as exogenously fixed.

Let r denote the subjective rate of time preference and let U^j and E be the expected present values of being unemployed, $j = e, s$, and employed, respectively. The value functions can then be written as:

$$\begin{aligned} rU_i^e &= \max_{s_i^e} \left\{ \upsilon(B, s_i^e) + s_i^e \alpha(\theta) \left(E - U_i^e \right) - \mu\pi \left(s_i^e \right) \left(U_i^e - U^s \right) \right\} && (5) \\ rU_i^s &= \max_{s_i^s} \left\{ \upsilon(Z, s_i^s) + s_i^s \alpha(\theta)(E - U_i^s) \right\} && (6) \\ rE_i &= \upsilon(w_i, \overline{h}) - \phi(E_i - U^e) && (7) \end{aligned}$$

The unemployed worker chooses search effort to maximize rU_i^j. The first-order conditions are given by:

$$\upsilon_s(B, s^e) + \alpha(\theta)(E - U^e) - \mu\pi_s \left(s^e \right) \left(U^e - U^s \right) = 0 \quad (8)$$

$$\upsilon_s(Z, s^s) + \alpha(\theta)(E - U^s) = 0 \quad (9)$$

where partial derivatives with respect to search effort are indicated by subscript s. In these expressions we have imposed symmetry, i.e., we have made use of the fact that workers are identical and choose the same search effort. The first-order conditions convey the usual message:[6] at the optimum,

[6]The second-order conditions for a maximum are fulfilled by the concavity of $\upsilon(\cdot)$ and the linearity of π.

the marginal cost of search should equal the marginal benefits. The marginal cost is captured by foregone leisure, i.e., $\upsilon_s(B, s^e)$ and $\upsilon_s(Z, s^s)$. The marginal benefit involves the gain in utility associated with a transition to employment, i.e., $\alpha(\theta)(E - U^j)$, $j = e, s$. For the eligible worker, there is an additional benefit of more intensive search, as revealed by the third term on the right-hand side of (8). More intensive search reduces the probability of being sanctioned, thus prolonging the expected duration of benefit payments. This does not imply, however, that eligible workers necessarily search harder than sanctioned workers. The effect pulling in the opposite direction is $B > Z$: sanctioned workers gain more from finding a job than eligible workers since $E - U^s > E - U^e$ holds in equilibrium. Which effect dominates depends on the parameters of the UI system.

For the remainder of the paper, we assume that the instantaneous utility functions take the form:

$$\upsilon(m, l) = \ln m + \Gamma(l), \quad m = \{w, B, Z\},\ l = \{1 - \bar{h}, 1 - s^e, 1 - s^s\}$$

where m denotes (real) income, which depends on the worker's labor market position and $\Gamma(l)$ represents the value of leisure with $\Gamma'(l) > 0$ and $\Gamma''(l) < 0$. The employed worker receives a wage w; the eligible unemployed worker receives unemployment insurance, B; and an unemployed worker who has been exposed to a sanction receives unemployment assistance, Z. Note that, given the monitoring technology, this functional form implies that $Z = 0$ can never be optimal from a social point of view. Hence, in any welfare maximizing system unemployment assistance will be strictly positive.

2.4 Firms and Wage Bargaining

Assume that government expenditure on benefits and monitoring is financed by a proportional payroll tax paid by firms. Labor productivity is constant and denoted y. The cost of holding a vacancy is ky, with $k > 0$. Let V denote the present value of a vacant job and J the present value of an occupied job. The value functions are of the usual form:

$$rV = -ky + q(\theta)(J - V) \tag{10}$$

$$rJ = y - w(1 + t) - \phi(J - V) \tag{11}$$

where t is the proportional payroll tax rate. With free entry of new vacancies, $V = 0$, we obtain the wage cost as proportional to the marginal product of labor, i.e.,

$$w\,(1+t) = [1 - (r+\phi)k/q(\theta)]y \tag{12}$$

Defining $\omega \equiv w(1+t)$ and writing the right-hand side of this equation as $d\,(\theta)\,y$, we refer to $\omega = d(\theta)y$ as the zero profit condition, with $d'(\theta) < 0$.

The outcome of the Nash bargain

$$\max_{w_i} \left[E(w_i) - U^e\right]^{\beta} \left[J(w_i) - V\right]^{1-\beta}, \quad \beta \in (0,1)$$

is a relationship of the form:

$$\frac{E - U^e}{w v_w} = \frac{\beta}{1-\beta}\frac{J}{\omega} \tag{13}$$

where $V = 0$ and symmetry have been imposed. The Nash bargain implies a wage-setting relationship, i.e., a relationship between bargained wages and labor market tightness. We assume that the government fixes the replacement rates in this economy. Hence $Z = zw$ and $B = bw$ where z and b are policy parameters. The replacement rates are defined with respect to the economy-wide average wage which the individual employee perceives to be independent of his wage demands; therefore $\partial U^e/\partial w = 0$. Finally, the relative size of the benefit sanction is denoted by p, i.e. p satisfies $z = (1-p)b$.

2.5 Equilibrium

Our assumptions imply that the model has a convenient recursive structure; the model in Fredriksson and Holmlund (2001) has a similar structure. The zero-profit condition and the wage-setting relationship determine θ and ω. To see this, note that with free entry of vacancies we have $J = ky/q(\theta)$ and $\omega = d(\theta)y$, which imply that the right-hand side of (13) is increasing in θ but independent of s^j. Moreover, the left-hand side of (13) is a function of θ but independent of w given our chosen utility function and the fact that income during unemployment is proportional to the aggregate wage. It can also be shown that as long as search is optimally chosen $E - U^e$ does not depend on s^j. With θ determined, we get s^j from (8) and (9), since the differences in present values are independent of w. With θ and s^j determined, we obtain u^j and n from (1)-(3).

Notice that θ, ω, s^j, u^j and n are independent of the tax rate, t. The latter can be determined residually from the government's budget restriction, noting that the government uses the wage tax to finance benefits and monitoring costs:

$$twn = u^e bw + u^s zw + c(\sigma)\mu u^e w \tag{14}$$

With the tax rate determined, the worker's take-home wage is obtained from $w = \omega/(1+t)$.

3 Optimal Unemployment Insurance

The optimal unemployment insurance system involves four instruments: b, p, μ, and σ. We assume that the government has a utilitarian objective function (W). We ignore discounting (see section 4.3 for the case with $r > 0$); hence it is valid to compare alternative steady states without considering the adjustment process. No discounting also has the convenient implication that the steady state flow of profits is zero.[7] Thus, the welfare objective is given by $W = u^e rU^e + u^s rU^s + nrE$, which simplifies to an employment-weighted average of instantaneous utilities with $r \to 0$:

$$W = nv(w, \overline{h}) + u^e v(B, s^e) + u^s v(Z, s^s) \tag{15}$$

The optimal policy maximizes (15) subject to the market equilibrium conditions, $s^j = s^j(b, p, \mu, \sigma)$ and $\theta = \theta(b, p, \mu, \sigma)$, as well as the balanced budget constraint, $t = t(b, p, \mu, \sigma)$. Let $\rho = \{b, p, \mu, \sigma\}$ denote the vector of policy parameters. Hence the vector of first-order conditions is given by $(dW/d\rho) = 0$.

Before proceeding to the numerical results it is useful to state two analytical results. First of all, the key result in Fredriksson and Holmlund (2001) applies directly. The following proposition reiterates proposition 2 in Fredriksson and Holmlund (2001).

Proposition 1 *The optimal policy involves $p > 0$.*

Proof. The proof is by contradiction. First, note that an optimal replacement rate is strictly positive, $b > 0$. This follows from the utility function

[7] Write the aggregate flow of profits as $\pi = n(y - w(1+t)) - kyv$, use eq. (12) and the flow equilibrium condition $q(\theta)v = \phi n$ and obtain $\pi = nrky/q(\theta)$.

above with $\ln B = \ln(bw)$. Consider the trial solution $p = 0$. At $p = 0$, the first-order condition for σ has a solution at $\sigma = 0$ because $c'(\sigma) \geq 0$. Moreover, the condition for μ is irrelevant. So, let us fix μ at some arbitrary, but interior, value: $\mu^0 \in (0, \infty)$. The uniform benefit structure ($p = 0$) cannot be optimal if $(dW/dp) > 0$ at $p = 0$. Some manipulations of the first-order condition for p using the optimality of b (i.e. $dW/db = 0$) yields

$$\left(\frac{dW}{dp}\right)_{p=0} = \frac{\partial W}{\partial s^s}\frac{\partial s^s}{\partial p} > 0 \tag{16}$$

where $\partial W/\partial s^s > 0$ denotes the partial derivative of welfare with respect to s^s holding θ constant, and $\partial s^s/\partial p > 0$ is also defined holding θ constant. ■

There are two key mechanisms that yield the sign of (16): there is a taxation externality associated with search and there is an "entitlement effect". The taxation externality derives from the fact that, given that some insurance is optimal ($b > 0$), taxes are required to finance unemployment expenditure. Individuals, however, do not take into account that taxes can be lowered if search intensity (and hence employment) increases. Therefore, $\partial W/\partial s^s > 0$. Moreover, the so called entitlement effect (c.f. Mortensen, 1977) will operate in this setting. Increasing the penalty will be conducive to search among those who are sanctioned since individuals will be eager to find a new job in order to qualify for (to be *entitled* to) UI benefit receipt, $\partial s^s/\partial p > 0$. As a corollary to proposition 1, the optimal policy will involve an interior μ. In other words, the two tiered benefit structure, $b > 0$, $p > 0$, and $\mu \in (0, \infty)$, dominates the uniform benefit structure in welfare terms.

Another interesting question is whether it will be optimal to have the sanctioning rate depend on search intensity, given an optimal choice of b, p, and μ. Since the inspection of search is the defining characteristic of the monitoring and sanctions system in this setting, we can equally well phrase the question as: Given an optimal choice of a UI system with time limits, is it optimal to introduce a system of monitoring and sanctions? The following proposition gives the condition when the answer turns out to be affirmative

Proposition 2 *Let $\hat{\rho} = \{b, p, \mu, \sigma = 0\}$ denote the solution to the restricted problem of optimal UI design. Then, the optimal policy will involve $\sigma > 0$ if*

$$\left(\frac{b}{\mu}\frac{\lambda + (1-\lambda)(1-p)}{\lambda s^e + (1-\lambda)s^s}\frac{\partial s^e}{\partial \sigma}\right)_{\rho=\hat{\rho}} > c'(0)$$

Proof. The proof proceeds as follows. Given that the two-tiered benefit structure is optimal, there are interior solutions to the first-order conditions $(dW/db) = 0$, $(dW/dp) = 0$, and $(dW/d\mu) = 0$. An UI system with monitoring and sanctions must be optimal if $(dW/d\sigma) > 0$ at the point where $\sigma = 0$ and the remaining first-order conditions hold. Some manipulations of the first-order condition for σ using $(dW/d\mu) = 0$ yields

$$\frac{dW}{d\sigma} = \frac{\partial W}{\partial s^e}\frac{\partial s^e}{\partial \sigma} - \frac{c'(0)\mu u^e}{(1+t)n} \tag{17}$$

where $\partial W/\partial s^e > 0$ denotes the partial derivative of welfare with respect to s^e holding θ constant, and $\partial s^e/\partial\sigma > 0$ is also defined holding θ constant. Introducing the explicit expression for $\partial W/\partial s^e$ and rewriting slightly, we get

$$\text{sign}\left\{\frac{dW}{d\sigma}_{\rho=\hat{\rho}}\right\} = \text{sign}\left\{\left(\frac{b}{\mu}\frac{\lambda + (1-\lambda)(1-p)}{\lambda s^e + (1-\lambda)s^s}\frac{\partial s^e}{\partial \sigma}\right)_{\rho=\hat{\rho}} - c'(0)\right\}$$

■

Equation (17) illustrates the basic trade-off in introducing a monitoring and sanctions system. A monitoring and sanctions system restores the search incentives among the eligible, $\partial s^e/\partial\sigma > 0$. Again, this is a good thing since there is a taxation externality which is not taken into account in the private determination of search. However, inspecting search consumes real resources as indicated by the second term in (17). If this cost is sufficiently high, the monitoring and sanctions system will not be introduced.[8]

Proposition 2 relates to the result in Boone and van Ours (2000). Their key result is that a monitoring and sanction system will be more efficient in restoring search incentives than overall benefit reductions. This result is derived by means of numerical solutions to a model which is essentially identical to the present one, but with $c' = 0$. Proposition 2 shows that their conclusion holds analytically. In addition it extends their result further: given $c' = 0$, a system with monitoring and sanctions will dominate the two-tiered benefit system analyzed by Fredriksson and Holmlund (2001).

By inspection of (16) and (17), the extent that search responds to incentives is going to be crucial for the amount of benefit differentiation and the argument for introducing monitoring and sanctions.

[8]If introducing a sanction system involves a fixed set-up cost (besides C), then clearly the set-up cost should not be too big either.

4 Numerical Analysis

We have calibrated the model numerically so as to provide some information on plausible numbers. The basic time unit is taken to be a quarter and the matching function is Cobb-Douglas, $H = aS^{1-\eta}v^{\eta}$, where we set $\eta = 0.5$.[9] We fix hours of work exogenously to $\overline{h} = 0.75$ and use the following parametrization of the value of leisure:

$$\Gamma(l) = \chi \frac{l^{\kappa} - 1}{\kappa} \tag{18}$$

where $\kappa < 1$. The marginal product of labor is normalized to unity and we impose $\beta = \eta$. This assumption corresponds to the Hosios (1990) efficiency condition for the case where agents are risk neutral.[10]

We calibrate the model for a uniform benefit system ($p = 0$) with a replacement rate of $b = 0.3$. The parameters a and χ are chosen with an eye towards vacancy duration and search intensity. We set $a = 1.7$ and $\chi = 0.6$. Remaining parameters (k, κ, and ϕ) are calibrated such that expected unemployment duration is one quarter, the partial equilibrium elasticity of the job hazard with respect to unemployment benefits equals -0.5, and the unemployment rate equals 6.5 percent. The calibrated values imply, e.g., that the inflow into unemployment is 28 percent a year and that the expected vacancy cost is almost a quarter of production. In the baseline calibration, the expected vacancy duration is close to half a quarter and search intensity equals $s = 0.7$.

Table 1 summarizes the parameter values in the baseline economy.

Table 1: Baseline parameters

Interest rate (= rate of time preference)	$r = 0$
Job destruction rate	$\phi = 0.069519$
Leisure value	$\kappa = 0.239419$, $\chi = 0.6$
Matching function	$\eta = 0.5$, $a = 1.7$
Wage negotiations	$\beta = \eta = 0.5$
Production	$y = 1$
Vacancy costs	$k = 1.98335$

[9] Broersma and Van Ours (1999) give an overview of recent empirical studies of the matching function. They find that a value of η of 0.5 is a reasonable approximation.

[10] Note that in our case workers are risk averse and hence $\beta = \eta$ is not sufficient for efficiency.

We also calibrate an alternative "less flexible" economy which has an identical unemployment rate but search is less responsive to incentives.[11] We obtain this characterization by lowering the constant in the matching function by 15 percent to $a = 1.445$ and compensating for this by a reduction in χ. A reduction in χ means that individuals place a lower value on leisure. The consequences of this are twofold: first, they are willing to search harder; second, and crucially, search is less responsive to changes in incentives. In particular, the partial equilibrium elasticity of the job hazard with respect to unemployment benefits equals -0.2 in this case. This is at the lower end of the interval given by Layard et al. (1991). The value of χ implying an unemployment rate of 6.5 percent, given the reduction in a, is $\chi = 0.364165$. The key outcomes in the base runs are reported in detail in columns 1 and 4 in Table 2.

4.1 Infinite vs Finite UI Benefit Duration

There are two natural focal points in the model. The first is the optimal uniform system (which has infinite UI duration: $\mu = 0$); the second is a system with optimal time limits (finite UI duration: $\mu > 0$ but $\sigma = 0$).

The last line of Table 2 presents welfare gains associated with particular policies. The welfare gain has the interpretation of a "consumption tax" (in percent) that equalizes welfare across two policy regimes. To be specific, let W^R represent welfare associated with a reference policy regime and W^A welfare associated with an alternative policy. Our measure of the welfare gain of policy A relative to policy R is given by the value of the tax rate τ that solves $W^A[(1-\tau)m;\cdot] = W^R$. With logarithmic utility functions we have $\Delta W \equiv W^A - W^R = -\ln(1-\tau) \approx \tau$. The welfare gains are always reported relative to the optimal uniform system. In order to compare, say, the system with time limits with the base run, one only has to take the difference between the two entries for the welfare gain (ΔW).

[11] Let us be clear here: the key is that search intensity in the "less flexible" economy is less elastic than search in the baseline economy. We coin this economy "less flexible" for want of a better word.

Table 2: Numerical results without monitoring and sanctions

	Baseline economy			Less flexible economy		
	Base run	Optimal uniform	Optimal time limit	Base run	Optimal uniform	Optimal time limit
	(1)	(2)	(3)	(4)	(5)	(6)
b	0.300	0.363	0.553	0.300	0.441	0.557
p	0	0	0.410	0	0	0.305
μ	–	–	1.556	–	–	1.199
σ	–	–	0	–	–	0
s^e	0.700	0.607	0.556	0.839	0.734	0.715
s^s	–	–	0.670	–	–	0.774
θ	0.705	0.578	0.528	0.680	0.470	0.446
u (%)	6.50	8.13	8.19	6.50	8.73	8.75
u^e (%)	6.50	8.13	2.84	6.50	8.73	3.36
u^s (%)	–	–	5.35	–	–	5.39
w	0.913	0.909	0.908	0.903	0.897	0.896
t (%)	2.09	3.21	3.61	2.09	4.22	4.70
ΔW (%)	–0.33	0	0.21	–0.90	0	0.11

In columns 2 and 5 of Table 2, we report the results of determining the optimal uniform replacement rate. The optimal replacement rate in the baseline economy is around 36 percent. A higher replacement rate reduces search incentives and incentives for wage restraint, so unemployment increases. With the optimal uniform replacement rate, unemployment rises to 8.1 percent. Individuals living in the baseline economy would be willing to pay 0.33 percent of consumption to move from a replacement rate of 30 percent to an optimal uniform one. The optimal uniform replacement rate in the "less flexible" economy is higher since the cost of raising the replacement rate in terms of reducing search incentives is lower. The replacement rate equals 44 percent in the less flexible economy and unemployment increases to 8.7 percent. Individuals in the less flexible economy would be willing to pay 0.9 percent of consumption in order to live in the optimal uniform system.[12]

The characteristics of the optimal system with time limits are given in column 3 and 6. In the baseline economy, benefit differentiation is substantial and the duration of UI benefit receipt is fairly short – the value of μ translates

[12]Notice that one should not compare the values of the consumption taxes across the two economies since the utility functions are different.

to an expected duration of around two months. The UI replacement rate amounts to 55 percent of the wage; the penalty associated with the loss of entitlement is around 41 percent. The benefit system with limited duration is substantially more generous than the system with infinite duration; with finite duration, unemployment expenditure per non-employed equals 40.5 percent. When search is less elastic, the UI replacement rate is about the same (58 percent) as in our base case. However, the penalty associated with loosing entitlement is decidedly smaller (31 percent), and the expected duration of UI receipt is longer (around 11 weeks). The unemployment rate is only marginally higher than in the uniform system.

Because the government has two additional instruments (μ and p) besides b it is not surprising that the welfare gain in the exogenous time limit case exceeds the welfare gain in the optimal uniform case in both economies. The relative gain of introducing time limits is, however, smaller in the less flexible economy than in the baseline economy. Also note that unemployment goes up by moving from the optimal uniform system to exogenous time limits. In other words, unemployment is not a sufficient statistic for welfare in this case.

4.2 Monitoring and Sanctions

This section evaluates the case for monitoring and sanctions and calculate the optimal monitoring and sanctions system. We also discuss the trade-off between monitoring and sanctions and investigate whether the penalties and sanctioning rates generated by the model are in broad conformity with the data.

4.2.1 Are Monitoring and Sanctions Optimal?

The argument in favor of monitoring and sanctions hinges crucially on the costs of this system. Unfortunately, the cost associated with monitoring and sanctions is something of a black box. Therefore, we give an upper bound on the marginal cost below which monitoring and sanctions are an ingredient of the optimal system. Since this upper bound turns out to be very high, we go on to characterize the optimal UI system with monitoring and sanctions.

Is it optimal to introduce monitoring and sanctions? In proposition 2 we stated the condition when the introduction of monitoring and sanctions represents a welfare improvement. For the introduction of monitoring and sanctions to be a welfare improvement relative to the case with time limits, $c'(0)$ has to be less than the gain as represented by greater search incentives among UI recipients with monitoring. We have calculated the cut-off value for our two economies. In our base case, this cut-off value ($\hat{c}$) equals $\hat{c} = 0.076$; in the alternative case, we have $\hat{c} = 0.047$. Both of these numbers have to be considered extremely high. Since the marginal product of labor and the labor force are normalized to unity, we can relate these cut-off values to (private sector) GDP by dividing by the employment rate (which is around 92 percent in the optimal system with time limits). So, the calculated cut-off values suggest that as long as the marginal cost is no greater than $4.7/0.92 = 5.1$ ($7.6/0.92 = 8.3$) percent of GDP, it is optimal to introduce monitoring and sanctions. Since these numbers are very large, the introduction of monitoring and sanctions is most likely a welfare improvement relative to the case with time limits.

What is the optimal design of a monitoring and sanctions system? This clearly depends on the exact form of the cost function $c(\sigma)$. Assume that $c(\sigma)$ takes the form of $c(\sigma) = \delta\sigma$. To estimate a reasonable value for δ, we used Swedish data on the relative number of employees at the Public Employment Service (PES), since PES officers are responsible for monitoring job search in Sweden. We also used information on how often each PES employee meets a particular unemployed, and the fraction of total time that the PES officer spends in meetings with the unemployed. This calculation, which is presented in greater detail in Appendix B, suggests that the marginal cost of monitoring is in the order of $c(\sigma) = \delta\sigma = 0.00785$. Provided that $\sigma \geq 0.785$ in Sweden, then $\delta = 0.01$ is a conservative estimate. We also conduct an alternative calculation where $\delta = 0.02$. Note that in both cases $\delta < \hat{c}$ and hence monitoring and sanctions improve welfare.

Table 3: Numerical results with monitoring and sanctions

	Baseline economy		Less flexible economy	
	$\delta = 0.01$	$\delta = 0.02$	$\delta = 0.01$	$\delta = 0.02$
	(1)	(2)	(3)	(4)
b	0.626	0.617	0.619	0.610
p	0.564	0.584	0.511	0.570
μ	1.207	1.039	1.017	0.757
σ	1	1	1	1
s^e	0.755	0.746	0.856	0.848
s^s	0.737	0.753	0.842	0.863
θ	0.407	0.408	0.371	0.373
u (%)	7.87	7.88	8.47	8.48
u^e (%)	5.75	5.96	7.07	7.37
u^s (%)	2.12	1.92	1.40	1.11
w	0.906	0.906	0.894	0.894
t (%)	4.60	4.66	5.33	5.35
ΔW (%)	0.82	0.75	0.55	0.48

Table 3 presents some numbers that correspond to the optimal systems in each economy for the two values of δ. The optimal system involves $\sigma = 1$ given our assumption $\sigma \in [0, 1]$. This particular result should be taken with a due grain of salt given the uncertainty about the costs of monitoring and the properties of the inspection technology. It is nevertheless interesting to note that a system with monitoring and sanctions are associated with non-trivial welfare gains relative to the alternatives characterized in Table 2. Also note that the optimal UI replacement rate is higher when we introduce monitoring and sanctions. Both economies experience a slight fall in unemployment as compared to Table 2, a result driven by a substantial increase in search effort among the unemployed (particularly those eligible for UI who now face additional incentives to search). Finally, note that the fraction of unemployed with a sanction is considerably lower in Table 3 than in the columns with exogenous time limits in Table 2. Less people need to be penalized in a monitoring system in order to get similar welfare and search incentive effects.

Table 3 indicates that the trade-off between monitoring and sanctions depends on the costs of monitoring: the higher the cost, the lower the monitoring rate and the higher the penalty. We have examined this trade-off in greater detail. In particular we have calculated optimal combinations of μ and p for different values of δ, where δ is varied from zero to (implausibly) large numbers. We set $\sigma = 1$ and allow b to adjust optimally. In order to approach the Beckerian corner solution ($\mu \to 0$, $p \to 1$), monitoring costs need to be extremely high. For example, if $\delta = 0.14$ the optimal system in the baseline economy features $p = 0.929$ and $\mu = 0.234$. Risk aversion in combination with a random monitoring technology implies that it is generally not optimal to impose the maximal sanction.

4.2.2 A Brief Look at the Data

Having calculated the optimal systems with monitoring and sanctions it is tempting to relate the predictions of the model to the data. Some of the parameters of the monitoring and sanctions system are of course unobservable. However, there are observations on the UI replacement rates, the penalties for violating search requirements, and the associated sanctioning rates. Presumably, there is a lot of noise in the data pertaining to sanction rates. Nevertheless, there is great variation in these data as is clear from Grubb (2001). It seems that the US and Switzerland are the extreme cases in terms of having systems with a large number of sanctions. In the US in the late 1990s, around 10 percent of beneficiaries were sanctioned each quarter for behavior during the benefit period. In addition, some 25 percent of the (stock of) eligible unemployed were "sanctioned" because they exhausted their benefits.[13] Based on these data, the quarterly sanction rate in the US would be in the order of 35 percent. With the exception of Switzerland, sanctions during the benefit period are substantially less common in the European countries. In fact, the sanction rates are typically lower than one percent per quarter; see Grubb (2001) for further details.

The number of sanctions seems to be inversely related to the severeness of the penalty. In the US, the normal sanction for a job search infringement is a loss of benefits for one week.[14] In Sweden, on the other hand, the penalty until recently was the loss of benefits for twelve weeks.[15]

[13] This estimate is a crude average for the period 1995-2000. The number of exhaustions per quarter amounted to some 600 000 individuals, the number of unemployed to 6.5 millions, and the fraction eligible for UI to 35 percent. Source: US Department of Labor (labor force statistics and UI program statistics).

[14] Notice, though, that there is a rather harsh "penalty" associated with the expiration of UI benefits in the US. In 1991, benefits were reduced by more than 60 percent when benefits expired and the individual was forced to claim welfare benefits instead; see Wang and Williamson (1996).

[15] The Swedish system has recently been changed in the direction of smaller penalties.

What does the model have to say about the number of sanctions? Figure 1 addresses this question by plotting the sanctioning rates against σ and assuming $\delta = 0.01$. In addition to the baseline and the less flexible economy, we also consider an economy with low turnover.[16] Sanctioning rates decline in σ for two reasons: firstly, for given s^e, a rise in σ reduces $\pi(s^e)$; and, secondly, a rise in σ raises s^e.

Table 4: Quarterly sanction rates according to the model

	Baseline economy	Less flexible economy	Low turnover economy
$\delta = 0.01$	0.296	0.146	0.242
$\delta = 0.02$	0.264	0.115	0.221

When $\sigma = 1$, as is optimal given our assumptions, the quarterly sanction rates hover between 10 and 30 percent depending on the exact assumptions; see Table 4. The number of sanctions in the baseline economy best conform to sanctioning data for the US. To get at the numbers for the typical European country, it appears that one would have to apply a combination of less elastic search, lower turnover, and higher monitoring costs.

4.3 Discounting and the Distributional Effects of UI

Discounting ($r > 0$) adds another dimension to our analysis of optimal UI policies. With discounting it makes a difference whether you are currently unemployed with or without a sanction. Hence it is relevant to analyze the distributional effects of optimal UI policies. Which category of workers are the prime gainers from the policies that we are considering?

Taking discounting into account also adds a complication to the analysis.[17] With discounting it is no longer the case that the aggregate profits of the firms are zero. Hence, we have to take a stance on what happens with these profits. It turns out that our results below do not depend much on

[16] The "low turnover" economy has a lower job destruction rate ϕ (around 22 percent per year) and higher value for a in the matching function to keep unemployment at the baseline value of 6.5 percent.

[17] Moreover, it is no longer possible to show analytically that the declining benefit sequence is to be preferred over the flat benefit sequence; see Fredriksson and Holmlund (2001) for further details. This is due to the wage pressure effect emphasized by Cahuc and Lehmann (2000). However, in our numerical examples the case for restoring search incentives is always stronger than the wage pressure effect.

what choices we make. Again, the objective function for the social planner is steady state welfare

$$W = nrE + u^e rU^e + u^s rU^s$$

and we assume that profits from firms are used to finance the UI system. The government's budget constraint is thus modified to

$$twn + nrky/q(\theta) = u^e bw + u^s zw + c(\sigma)\mu u^e w$$

where $nrky/q(\theta)$ is the aggregate profit. In this way we do not need to worry about the distribution of the profits over the three categories of workers.[18]

Table 5 presents the numerical results for the three different UI systems considered. We conduct the calculations for the baseline economy and report the results for two values of the discount rate: 5 and 10 percent on an annual basis. The marginal cost of monitoring (δ) is set equal to $\delta = 0.01$.

We first note that discounting only marginally affects the actual policies chosen. The parameters of the UI systems are very similar to the values reported in Tables 2 and 3. Consequently, the outcomes in terms of search intensities and unemployment are close to the ones we have reported earlier.

[18]Note that this assumption is approximately equivalent to assuming that workers own firms in proportion to their income or that there is a separate group of risk neutral firm owners in the economy.

Table 5: Numerical results with discounting in the baseline economy

	Annual discount rate: 0.05			Annual discount rate: 0.10		
	Uniform system	Time limit	Monitoring and sanctions	Uniform system	Time limit	Monitoring and sanctions
	(1)	(2)	(3)	(4)	(5)	(6)
b	0.358	0.544	0.620	0.353	0.536	0.615
p	0	0.409	0.570	0	0.408	0.576
μ	–	1.565	1.213	–	1.574	1.221
σ	–	0	1	–	0	1
s^e	0.609	0.557	0.757	0.610	0.560	0.760
s^s	–	0.669	0.738	–	0.670	0.739
θ	0.573	0.524	0.404	0.567	0.520	0.401
u (%)	8.15	8.21	7.88	8.18	8.23	7.89
u^e (%)	8.15	2.83	5.75	8.18	2.82	5.76
u^s (%)	–	5.38	2.12	–	5.41	2.12
w	0.910	0.909	0.907	0.910	0.910	0.908
t (%)	1.96	2.40	3.54	0.73	1.20	2.49
ΔW (%)	0	0.21	0.83	0	0.20	0.84
ΔrE (%)	0	0.22	0.83	0	0.22	0.84
ΔrU^e (%)	0	0.27	1.03	0	0.34	1.25
ΔrU^s (%)	0	−0.001	0.10	0	−0.22	−0.63
$\Delta r\tilde{U}$ (%)	0	0.09	0.78	0	−0.03	0.75

Notes: The welfare gains are measured relative to the optimal uniform systems. The average welfare change for unemployed workers is denoted $\Delta r\tilde{U}$, where $r\tilde{U} =$ $r\tilde{U} = (u^e/u)rU^e + (u^s/u)rU^s$. The marginal cost of monitoring equals $\delta = 0.01$

The most interesting information provided by this exercise is reported at the bottom end of Table 5. The last four lines show the welfare gains for different groups of individuals relative to the optimal uniform system. The ranking of the welfare gains for the three groups conforms to what one would expect. The eligible unemployed are the prime gainers, while those on a sanction are the ones that gain the least from introducing time limits or monitoring and sanctions. However, it is interesting to see that a monitoring and sanctions system makes everyone better off compared to a system featuring a pure time limit. Unlike the time limit system, the monitoring and sanctions system always improve the welfare of an average unemployed worker. Moreover, monitoring and sanctions even increases the welfare for those on a sanction for the moderate (and, in our view, more reasonable) annual discount rate of 5 percent.

The reason why a monitoring system makes even the unemployed with a sanction better off than in a uniform system is the entitlement effect. Although the unemployed with a sanction have a lower per period utility than the unemployed under a uniform system, they can look forward to a more generous unemployment insurance scheme after they lose their next job.[19] When the discount rate is raised to 10 percent, the relative valuation of the present state increases. In this case, the entitlement effect is not sufficient to compensate for the direct utility loss.

5 Concluding Remarks

In this paper we have analyzed the design of optimal unemployment insurance in a search equilibrium framework where search effort among the unemployed is not perfectly observable. We have examined to what extent the optimal policy should involve monitoring of search effort and benefit sanctions if observed search is found insufficient. The results suggest that the introduction of a system with monitoring and sanctions represents a welfare improvement for reasonable values of the monitoring costs. Those costs would have to be implausibly high – higher than five percent of GDP – for this conclusion not to hold.

The policy prescription following from our analysis is thus different from Becker's (1968) well known result, where the penalty should be maximal and the probability of getting caught should be close to zero. There are two key assumptions delivering our results. First, individuals are risk averse and, second, monitoring is imperfect. With imperfect monitoring some individuals will be sanctioned even though they search to rule and giving them the maximal penalty is not optimal with risk aversion.

While we are reasonably comfortable in saying that monitoring and sanctions represent a welfare improvement, it is much more difficult to give clear advice on the characteristics of such a system. The reason for this conclusion is that the exact formulation of the monitoring and sanctions system depends on the cost of running such a system. Unfortunately, the cost of running the system is something of a black box.

[19] The value of UI is higher with monitoring than with time limits for two reasons. First, the UI replacement rate is higher. Second, the rate of being transferred to the "sanctioned state" is a lot lower with monitoring.

An issue that we have not addressed is the possibility that formal search requirements may induce individuals to use formal rather than informal search methods and therefore bring little increase in total search intensity. Nevertheless, it is likely that *general* search requirements – such as the number of job applications filed during a week – should minimize the risk of substitution between search channels. Presumably, substitution is going to be more severe in systems where search requirements are linked to formal channels such as referrals by the public employment service. On this account, search requirements specified in terms of independent job search, as used in the US, the Netherlands and Switzerland, seem to be preferable.

Appendix A: The Sanctioning Probability

This appendix addresses the "structural" interpretation of our sanctioning probability: $\pi(s^e) = 1 - \sigma s^e$. Suppose, realistically, that benefit administrators observe search with error: $s_o^e = s^e + \varepsilon, \varepsilon \in [\varepsilon_L, \varepsilon_U]$. Since $s^e \in [0,1]$ then so should s_o^e. This in turn implies restrictions on $\varepsilon_L, \varepsilon_U$. If $s_o^e \in [0,1]$, it must be true that $\varepsilon_L = -s^e$ and $\varepsilon_U = 1 - s^e$.

Let us introduce a parameter that indexes the extent of observation error. In particular let $\varepsilon \in [-(1-\tilde{\sigma})s^e, (1-\tilde{\sigma})(1-s^e)]$. If $\tilde{\sigma} = 1$, there is no observation error. If $\tilde{\sigma} = 0$, observed search belongs to the entire admissible range. Suppose also that ε is uniform. Then $\tilde{\sigma} = 0$ is a completely random inspection technology. We think of $\tilde{\sigma}$ as a parameter that the central government can invest resources in improving.

An individual is sanctioned whenever $s_o^e \leq R$, where $R \in [0,1]$ denotes the search requirement. The probability of being sanctioned given that the individual supplies s^e units of search is then

$$\pi = \int_{\varepsilon_L}^{R-s^e} \frac{1}{\varepsilon_U - \varepsilon_L} d\varepsilon = \frac{R - s^e}{\varepsilon_U - \varepsilon_L} - \frac{\varepsilon_L}{\varepsilon_U - \varepsilon_L}$$

Since $\varepsilon_U - \varepsilon_L = 1 - \tilde{\sigma}$ and $\varepsilon_L = -(1-\tilde{\sigma})s^e$, we get

$$\pi = \frac{R}{1-\tilde{\sigma}} - \frac{\tilde{\sigma}}{1-\tilde{\sigma}} s^e$$

Now we want to impose some restrictions on the parameters of the inspections technology $(R, \tilde{\sigma})$ to make sure that $\pi \in [0,1]$ for all $s^e \in [0,1]$. We impose the following conditions

1. If $s^e = 0$ then $\pi = 1$.

2. If $s^e = 1$ then $\pi \in [0, 1]$.

The first condition gives $R = 1 - \tilde{\sigma}$. Given $R = 1 - \tilde{\sigma}$, the second condition yields $\tilde{\sigma} \in [0, 0.5]$. The conditions we impose on the parameters thus imply that an individual who searches full time is sanctioned with positive probability, i.e., there is a probability of making Type II errors for all values of s^e.

In sum, the above assumptions lead to the following formulation for π

$$\pi = 1 - \frac{\tilde{\sigma}}{1 - \tilde{\sigma}} s^e, \; \tilde{\sigma} \in [0, 0.5]$$

or alternatively, defining $\sigma = \tilde{\sigma}/(1 - \tilde{\sigma})$

$$\pi = 1 - \sigma s^e, \; \sigma \in [0, 1]$$

which is what we have in the main text.

Appendix B: Estimating the Marginal Cost of Monitoring

To obtain a reasonable value for the cost of monitoring an additional individual ($c(\sigma)$) we performed the following calculation. We relied on data from Sweden, where PES administrators are responsible for monitoring whether unemployed individuals have searched to rule or not. Three sources of information were used: (*i*) the relative number of employees at the PES; (*ii*) the fraction of time that a PES officer meets with the unemployed; and (*iii*) the number of contacts between the PES officer and a particular unemployed individual. Information pertaining to items (*ii*) and (*iii*) is taken from Lundin (2000).

In the main text the total cost of the monitoring and sanctions system was specified as: $C = c(\sigma)\mu u^e w$. To get an approximate value for C we start by calculating the wage bill paid to individuals involved in monitoring. Since the labor force and the marginal product of labor are normalized to unity, the wage bill is measured relative to these items. The PES service employs approximately 10,000 individuals in Sweden, which translates to around 0.25 percent of the labor force. On average PES officers spend 30 percent of their time in meetings with the unemployed. Assuming that the unemployed are monitored each time they meet with a PES officer we have $C = 0.0025 \times 0.3w = 0.00075w$. Thus we have $C = c(\sigma)\mu u^e w = 0.00075w$. Turning to the left-hand side of this equation, we set the number of unemployed individuals eligible for UI to 5 percent. With this assumption, we only need an estimate of μ to get an estimate of $c(\sigma)$. The information used to estimate μ is derived from a question put to PES officers regarding the number of meetings with individuals searching for a job. When asked about their contact frequency, 35 percent of PES officers answered "at most once a month"; 34 percent answered "at most once every other month"; and 31 percent answered "at most once every quarter". Thus on average a PES officer has $(1 \times 0.35 + 0.5 \times 0.34 + 0.31/3) \times 3 = 1.91$ meetings with a particular unemployed per quarter. Hence we have $c(\sigma) = 0.00075/(\mu u^e) = 0.00075/(1.91 \times 0.05) \approx 0.00785$. There is still one unknown in this equation, however; the estimated value of $c(\sigma)$ pertains to a given value of σ. Assuming that $c(\sigma) = \delta\sigma$, we have $\delta = 0.01$ for $\sigma = 0.785$ and $\delta = 0.02$ for $\sigma = 0.785/2$.

References

Abbring, J.H., G.J. van den Berg and J.C. van Ours (1997) The effect of unemployment insurance sanctions on the transition rate from unemployment to employment, Working Paper, Tinbergen Institute, Amsterdam.

Becker, G. (1968) Crime and punishment: An economic approach, *Journal of Political Economy*, 76, 169-217.

Benus, J., J. Joesch, T. Johnson and D. Klepinger (1997) Evaluation of the Maryland unemployment insurance work search demonstration: Final report, Batelle Memorial Institute in association with Abt Associates Inc (http://wdr.doleta.gov/owsdrr/98-2/).

Black, D., J. Smith, M. Berger and B. Noel (2003) Is the threat of training more effective than training itself? Evidence from random assignments in the UI system, *American Economic Review*, 93, 4, 1313-1327.

Boone, J. and J. van Ours (2000) Modelling financial incentives to get unemployed back to work, Discussion Paper 2000-02, CentER for Economic Research, Tilburg University.

Broersma, L. and J.C. van Ours (1999) Job searchers, job matches and the elasticity of matching, *Labour Economics*, 6, 77-93.

Cahuc, P. and E. Lehmann (2000) Should unemployment benefits decreases with the unemployment spell?, *Journal of Public Economics*, 77, 135-153.

Fredriksson, P. and B. Holmlund (2001) Optimal unemployment insurance in search equilibrium, *Journal of Labor Economics,* 19, 370-399.

Garoupa, N. (1997) The theory of optimal law enforcement, *Journal of Economic Surveys*, 11 267-295.

Grubb, D. (2001) Eligibility criteria for unemployment benefits, in *Labour Market Policies and Public Employment Service*, OECD.

Hopenhayn, H.A. and J.P. Nicolini (1997) Optimal unemployment insurance, *Journal of Political Economy*, 105, 412-438.

Hosios, A.J. (1990) On the efficiency of matching and related models of search and unemployment, *Review of Economic Studies*, 57, 279-298.

Johnson, T. and D. Klepinger (1994) Experimental evidence on unemployment insurance work-search policies, *Journal or Human Resources*, 29, 695-717.

Lalive, R., J.C. van Ours and J. Zweimüller (2002) The effect of benefit sanction on the duration of unemployment, Discussion Paper 3311, Centre for Economic Policy Research.

Layard, R., S. Nickell and R. Jackman (1991) *Unemployment: Macroeconomics performance and the labour market*, Oxford: Oxford University Press.

Lundin, M. (2000) Tillämpningen av arbetslöshetsförsäkringens regelverk vid arbetsförmedlingarna, stencil 2000:1, Office of Labour Market Policy Evaluation (IFAU), Uppsala.

Mortensen, D.T. (1977) Unemployment insurance and job search decisions, *Industrial and Labor Relations Review*, 30, 505-517.

Pissarides, C.A. (1990) *Equilibrium Unemployment Theory*, Basil Blackwell, Oxford.

Polinsky, A.M. and S Shavell (1979), The Optimal Tradeoff between the Probability and Magnitude of Fines, *American Economic Review*, 69, 880-891.

Polinsky, A.M. and S. Shavell (2000) The economic theory of public enforcement of law, *Journal of Economic Literature*, 38, 45-76.

OECD (2001) *Labour Market Policies and Public Employment Service*, Paris.

Shavell, S. and L. Weiss (1979) The optimal payment of unemployment insurance benefits over time, *Journal of Political Economy*, 87, 1347-1363.

van den Berg, G.J. and B. van der Klaauw (2001) Counselling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment, Working Paper 2001:12, Institute for Labour Market Policy Evaluation (IFAU), Uppsala.

van den Berg, G.J., B. van der Klaauw and J.C. van Ours (2004) Punitive sanctions and the transition rate from welfare to work, *Journal of Labor Economics*, 22, 211-241.

Wang, C. and S. Williamson (1996) Unemployment Insurance with Moral Hazard in a Dynamic Economy, *Carnegie-Rochester Conference Series on Public Policy* 44, 1-41.

Figure 1: Quarterly sanction rates, δ=0.01